DOT/FAA/AM-95/13

Office of Aviation Medicine
Washington, D.C. 20591

Practical Color Vision Tests for Air Traffic Control Applicants: En Route Center and Terminal Facilities

DTIC
ELECTE
JUN 0 7 1995
S D
F

H.W. Mertens
N.J. Milburn
W.E. Collins

Civil Aeromedical Institute
Federal Aviation Administration
Oklahoma City, Oklahoma 73125

April 1995

Final Report

This document has been approved for public release and sale; its distribution is unlimited.

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

DTIC QUALITY INSPECTED 3

U.S. Department of Transportation

Federal Aviation Administration

19950605 007

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The United States Government assumes no liability for the contents or use thereof

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-95/13		

4. Title and Subtitle	5. Report Date
Practical Color Vision Tests for Air Traffic Control Applicants: En route Center and Terminal Facilities	April 1995
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Henry W. Mertens, Ph.D., Nelda J. Milburn, M. Ed., and William E. Collins, Ph.D.	
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
FAA Civil Aeromedical Institute P.O. Box 25082 Oklahoma City, OK 73125	
	11. Contract or Grant No.
12. Sponsoring Agency name and Address	13. Type of Report and Period Covered
FAA Office of Aviation Medicine Federal Aviation Administration 800 Independence Avenue, SW Washington, DC 20591	

15. Supplemental Notes

This work was performed under task AAM-B-94-HRR-190

16. Abstract

Two practical color vision tests were developed and validated for use in screening Air Traffic Control Specialist (ATCS) applicants for work at en route center or terminal facilities. The development of the tests involved careful reproduction/simulation of color-coded materials from the most demanding, safety critical color task performed in each type of facility. The tests were evaluated using 106 subjects with normal color vision and 85 with color vision deficiency. The en route center test, named the Flight Progress Strips Test (FPST), required the identification of critical red/black coding in computer printing and handwriting on flight progress strips. The terminal option test, named the Aviation Lights Test (ALT), simulated red/green/white aircraft lights that must be identified in night ATC tower operations. Color coding is a non-redundant source of safety-critical information in both tasks. The FPST was validated by direct comparison of responses to strip reproductions with responses to the original flight progress strips. Validity was high; Kappa=.91. The light point stimuli of the ALT were validated physically with a spectroradiometer. The test lights met the FAA and ICAO standards for colors of aircraft and aviation signal lights. The reliabilities of the FPST and ALT were estimated with Chronbach's alpha and were .93 and .98, respectively. The high job-relevance, validity, and reliability of these tests increases the effectiveness and fairness of ATCS color vision testing.

17. Key Words	18. Distribution Statement
Air Traffic Controllers Color Vision Standards Performance Tests Color Vision Tests	Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	15	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

PRACTICAL COLOR VISION TESTS FOR AIR TRAFFIC CONTROL APPLICANTS: EN ROUTE CENTER AND TERMINAL FACILITIES

INTRODUCTION

A medical color vision standard has been established for Air Traffic Control Specialists (ATCSs) by the U.S. Office of Personnel Management (1993) due to the use of color coding of critical information for several tasks, as verified by job-task analysis (Lahey, Veres III, Kuyk, Clark, & Smith, 1984). Previous research has also demonstrated that individuals with normal color vision rarely make errors in simulated ATCS color tasks, while many individuals with all levels of color vision deficiency frequently make errors (Adams & Tague, 1985; Kuyk, Veres, Lahey, & Clark, 1986, 1987; Mertens & Milburn, 1992a). Several clinical color vision tests have high predictive validity for identifying individuals who can perform particular ATCS color tasks without error (Mertens & Milburn, 1992b). Those tests in general tend to have low miss (false negative) rates, typically 5% or less. While all of the approved clinical tests had acceptable miss rates, false alarm rates tended to be somewhat higher, and were undesirably high with a few tests (Mertens & Milburn, 1992b). Some individuals who fail clinical tests may be able to perform ATCS color tasks as well as color vision normals. Practical color vision tests are, therefore, desirable for increasing the predictive efficiency and optimizing the validity of color vision testing for ATCS selection. The present research concerns the development and validation of practical, "work sample" color vision tests for secondary screening of ATCS applicants who are seeking employment at terminal and en route center air traffic control (ATC) facilities.

Optimally, color vision screening of ATCS applicants should provide both the ability to select all individuals who can perform safety-critical ATCS color tasks accurately, whether they have a color vision deficiency or not, and screen out all individuals who cannot perform those tasks accurately. A test with high predictive validity for performance in ATCS color tasks must have both a low miss (false negative) rate to ensure safety, and a low false alarm (false positive) rate to ensure fairness. The most direct way of accomplishing that objective is to create color vision screening tests that are "work sample" tests such that they accurately simulate the color-coded materials of the task and the viewing conditions of the work environment. The stimulus characteristics that affect color perception, such as color, color contrast, brightness, brightness contrast, and stimulus size, must be reproduced accurately. Viewing conditions that can affect color perception, such as intensity and color of ambient illumination and the viewing distance, must also be reproduced accurately. Such tests can be considered "practical tests" because they involve essentially the same materials and observation conditions that are involved in the actual work tasks (Jewell & Siegall, 1990). The practical tests should also: (1) simulate "normal" work situations, rather than more marginal observation conditions that sometimes occur, since those conditions may also degrade performance of individuals with normal color vision; and (2) involve very simple color identification and/or discrimination responses so that the tests assess only color vision ability. In addition to testing color vision appropriately for the job, the test's relevance to the job should be explained to the applicant and be sufficiently obvious so as to promote a positive response to the test situation.

Two practical color vision tests for en route center and terminal ATCS work options, that were developed in the present research, represent the most important, safety-critical tasks involving non-redundant color coding in each work situation. The practical tests were developed with the intention that they should be implemented for secondary color vision screening of ATCS applicants who fail clinical color vision testing during their initial pre-employment medical examina-

A-1

tion by an Aviation Medical Examiner. The two tests described here have subsequently been implemented, and the tests are administered at the office of the Regional Flight Surgeon in each of the nine Federal Aviation Administration (FAA) Regions. A practical test for a third ATCS work option involving work in Automated Flight Service Stations (AFSSs) is under development, but is not addressed in the present report.

Practical Color Vision Test for En Route Center ATCSs

The most demanding color vision task at en route centers involves accurately distinguishing between red and black computer printing and handwriting on Flight Progress Strips (FPSs). The test for this task is named the Flight Progress Strips Test (FPST). The FPSs are frequently referred to during an ATCSs time "on the boards" (Vortac, Manning, & Rotter, 1992). Each paper FPS is approximately 8 in wide by 1-5/16 in high, and is typically mounted in a plastic holder and stacked on a sloping bay located near the large radar cathode ray tube situation display of air traffic. There is at least one FPS for each aircraft that is controlled by the en route ATCS. The FPS contains a variety of information that is either computer printed by the ATC system's computers, or handwritten by the ATCSs. The FPS information describes the type of aircraft and the aircraft's situation, including altitude, departure, route, fixes, clearances, approach, and so forth. The most critical color coding involves the use of the colors black and red to distinguish between assigned versus non-assigned (planned) information concerning altitude, route, departure, approach, fixes, clearances, and so forth (FAA, 1989). Other uses of the black/red code in FPSs include differentiation between (1) coordination versus incorrect reporting altitude, (2) north versus south flights, and (3) east versus west flights.

The FPST requires the identification of red and black colors in reproductions of actual FPSs made by a photographic color separation process and lithographic reproduction process that is also used for color-critical fine art printing. The FPS copies were printed on the same 32 lb pale green ledger paper used by the FAA for FPSs. The FPST was validated by comparing color responses to the FPST with color responses to a test consisting of the original strips from which the copies were made. The criterion test containing actual strips was named the Criterion Flight Progress Strips Test (CFPST). The FPST and CFPST test materials were illuminated by a special light source developed to produce illumination of the color and intensity typical at actual ATCS work stations.

Practical Color Vision Test for Terminal ATCSs

The second test was designed for testing ATCS applicants who will work in air traffic control towers at terminal facilities, and who must, therefore, be able to identify and discriminate the color of aircraft lights at night. The color coding of aircraft lights is important for location of aircraft and determining their direction or orientation, in flight or on the ground, based on identification of red, green, and white position lights. The test is called the Aviation Lights Test (ALT) and it makes use of the body and mechanisms of the Farnsworth Lantern, a well-known clinical color vision test that involves identification of red, green, and white light points. Modification of the lantern's red and green filters was required to make the light points meet FAA and International Civil Aviation Organization (ICAO) color standards for aviation signal lights, but the brightness of simulated lights is similar to the Farnsworth Lantern. The administration and scoring procedures also differ from the Farnsworth Lantern as follows: (1) the ALT procedure involves a demonstration of each of the three test light colors prior to test trials; (2) the ALT always involves presenting three random series of the nine pairs of lights produced by the lantern, and requires scoring of responses to all 27 pairs of lights; and (3) the disposition criterion of the ALT is different and results in failure if two or more errors are made during the responses of 27 test trials; and (4) a very dim room illumination is used that approximates the dim illumination in the ATC tower cab at night. The ALT is a practical test because it involves actual red, green, and white point sized lights that accurately meet color specifications, and are within the range of intensities of aviation signal lights as seen at night from the ATC tower.

The basis of validity of the ALT is that the simulated lights actually meet the specifications for colored aviation signal lights as determined by colorimetric mea-

surements in the present research. The ALT test was given to individuals under both the design (dim) illumination, and under conditions of a normally lit office to determine the importance of variation in test room illumination in this study. It was predicted that test effectiveness would not be affected significantly by this variation in ambient illumination. That prediction is based on the general similarity of the ALT test stimuli to stimuli of the Farnsworth Lantern test, which has been shown to be insensitive to wide variation in ambient illumination (Dain, Honson, & Ang, 1988; Steen, Collins, & Lewis, 1974).

METHODS

Subjects

The tests were evaluated using 106 subjects with normal color vision and 85 with color vision deficiency. The ages of subjects with normal color vision ranged from 18 to 58 yr with a mean age of 31.2 yr and a standard deviation of 11.1 yr. The ages of subjects with color vision deficiency ranged from 18 to 61 yr with a mean age of 35.0 yr and a standard deviation of 12.5 yr. Normals included 63 men and 43 women. Color deficients included 84 men and 1 woman.

All subjects had at least 20/30 visual acuity in both near and distant vision, as measured by the Bausch and Lomb Orthorater vision tester. Subjects were recruited through advertisements in newsletters at local colleges and universities, and through local newspapers of the Oklahoma City (OK) metropolitan area. All subjects were paid an hourly wage.

Flight Progress Strips Test (FPST)

The FPST included a copy of each of 30 FPSs selected from many thousands of strips obtained from six en route centers; a day's output of strips was obtained from each center. The 30 original strips were used to make the CFPST. The FPSs were selected to be representative of the type fonts of the two types of computer printers that are used, the types of pens and pencils used for handwriting on strips, and representative of the range of darkness/lightness normally found in both computer printing and handwriting in the actual work environment.

Each group of computer-printed characters in the same line and column of an FPS, that did not contain a blank space and were the same color, were considered an item on the test section pertaining to computer printing. Each handwritten altitude in the altitude column was considered an item on the 21 strips (24 items) testing color vision for handwriting on strips. The FPS copies and instructions for the test were assembled in book form to make the FPST. Three of the 30 strips were used for demonstration and 27 served as test items. Subjects responded by identifying the colors red and black in computer printing on 6 strips containing 100 items and the color of handwriting on 21 strips containing 24 items. There was no time limit for responding and performance was assessed in terms of total number of errors and a pass-fail score. An error consisted in misidentifying the color of a red or black item on a strip. The failure criterion was two or more errors.

Criterion Flight Progress Strips Test (CFPST)

Performance of the CFPST served as the criterion for validation of the FPST. The CFPST was identical to the FPST, with the exception that it involved the actual FPSs, rather than the lithographic copies of those same FPSs that formed the FPST. The CFPST and FPST were essentially identical with regard to form of the test book, administration procedure, illumination, and scoring. The individual strips in each of the two parts of the FPST and CFPST, the computer printing and handwriting sections, were presented in different random orders in the two tests.

FPST Test Stand and Illumination

A test stand was constructed to hold the test books in a convenient testing position, and to provide consistent and appropriate illumination for both the CFPST and FPST. The stand is shown in Figure 1. This stand was designed to provide an illumination level of 59 lux at the center of the test page. That corresponded to the average workstation illumination as measured at two facilities, the Ft. Worth (TX) and Kansas City (KS) en route centers. This illumination level was based on measurements at four work stations at each facility. Measurements were taken in the FPS bay at the location

of the strip farthest from the task light for the bay, with the task light adjusted to the highest intensity. The task light used at en route centers for illuminating flight strip bays is a part of a standard console common to all en route centers; it contains a cool white fluorescent lamp. The present task light also used a cool white fluorescent lamp to make color characteristics of the test illumination realistic. The specific fluorescent tube used in the FPST light fixture was General Electric Part No. F4T5/CW. Each new tube was aged 50 hours before being used for testing and was tested to verify the 59 lux illumination level, with a tolerance of plus or minus 10 percent. Tubes were replaced after 200 h of usage. Measurements of illumination were made using a Photo Research Corporation (Chatsworth, CA) Fast Scan Spectroradiometer Model 702/703AM and reflectance standard Model RS-1. The 59 lux illumination level is comparable to illumination at other centers as reported by Adams and Tague (1985).

Aviation Lights Test

The test for applicants for the terminal work option, called the Aviation Lights Test (ALT), required the identification of the color of simulated color-coded aircraft lights in a dark environment as required for location and identification of aircraft and their direction in night air traffic control tower operations. The Aviation Lights Test made use of the body and mechanisms of the Farnsworth Lantern (Macbeth Corporation, Newburg, N.Y.), but modification of the Farnsworth Lantern's red and green lights was required to make the simulated lights meet FAA and ICAO specifications for aviation signal light colors [Department of Transportation (DOT)/FAA, 1971; ICAO, 1988]. White lights of the lantern met specifications. Appropriate custom filters were designed and manufactured by Kopp Glass Company (Pittsburgh, PA) and installed to modify the Farnsworth Lantern to produce the chromaticities which are plotted in Figure 1 along with the gamuts, or boundries, specified by the FAA and ICAO for aviation signal colors, red, green, and white. The new red and green filters were designed to make slight alterations in color, but maintain the same luminous intensity of stimuli, and therefore, retain the independence of color and intensity of light points.

The simulated aircraft lights were formed by two 2.5 mm circular apertures arrayed vertically and separated by 13.0 mm. The resulting size of light points at the 2.438 m (8 ft) viewing distance was 3.5 arcmin and their separation in the visual field of view between centers was 18.3 arcmin. The 18.3 arcmin separation would simulate the case of an aircraft with a 25 ft wing span at a distance of approximately 1,432 m (4,700 ft). The luminance of the surface of the small circular light sources in the lantern varied from 114.4 cd/m2 to 647.5 cd/m^2. That range of luminances of light points would approximately simulate aircraft position lights at distances of 190 m (623 ft) to 1,237 m (4058 ft). This assumes an aircraft position light luminous intensity of 40 cd for the shortest simulated distance and 300 cd for the longest simulated distance. Kaufman and Christensen (1987) state that the 40 to 300 cd range of position light intensities occurs among aircraft of different size in the U.S.; 40 cd is the minimum intensity required by FAR. There were 9 pairs of light points that simulated aviation lights. The 18 light points in the 9 pairs included 6 red, 6 green, and 6 white lights. The 9 pairs were presented in random order three times for a total of 27 trials. The misidentification of the color of either or both members of a pair constituted one error. The criterion for failure was two or more errors in 27 trials. That criterion was based on previous research with a similar task in which normal trichromats rarely made errors, and never more than one error (Mertens & Milburn, 1992a).

The illumination for the ALT test room, measured at the position of the ALT, is approximately 1 lux which is recommended as the upper limit for illumination at windows in the ATC tower (Illumination Engineering Society of North America, 1972). This nominal illumination level was achieved by placing a small desk lamp fixture with an clear incandescent 7 watt, 120 vac bulb approximately 3 ft behind the subject, with the lamp pointed toward the ceiling.

As mentioned in the introduction, the ALT test was also administered in a room with normal office illumination. The illumination at the observer's eye was 1065 lux when measured in a horizontal plane and 500 lux when measured in a vertical plane.

The data for the two conditions will be differentiated by calling the task with low, ATC tower illumination ALT-L, and the task with the high, office illumination ALT-H.

Diagnostic Color Vision Tests

The procedure used for classification of deficiencies involved the Nagel Type I anomaloscope to classify individuals with red-green color vision deficiency. Other tests were used to detect and diagnose the rare blue-yellow deficiencies that the Nagel anomaloscope does not detect, but none of the latter were found. These diagnostic methods were identical to those described by Mertens and Milburn (1992a).

Procedure

The Flight Progress Strips Tests, the Aviation Lights Test, and diagnostic tests were administered at four testing stations, each supervised by a trained laboratory technician. All anomaloscope testing was performed by the first author. The tests at each station took approximately 30-40 minutes to administer. The order of the two FPS tests was reversed for half of the subjects, as was the order of the two ALT tests. At least one other, different color vision test was given between the two FPS tests and the two ALT tests. The testing of each subject was performed in a three-hour test session. Within each session the four testing periods were separated by a 5-10 min break.

RESULTS AND DISCUSSION

Diagnostic Classification of Subjects

The number of subjects in each classification of type and degree of color vision deficiency are shown in Table 1. Representation of all categories of red-green color vision deficiency was obtained. The number of normal trichromats was 106 and the total number of deficients was 85. There were 35 deficients who were protans (pro) and 50 deficients who were deutans (deu). No individuals were found to have the blue-yellow type of color vision deficiency.

Relation of Performance on Flight Progress Strips Tests to Color Vision Classification

Pass/Fail Performance. The CFPST and FPST tests permit comparison of responses to actual flight progress strips and color copies of strips that comprise the practical FPST. The criterion for passing was "no more than one error" for the CFPST and the criterion was the same for the FPST. That criterion was based on use of the similar FPS materials in a previous experiment (Mertens & Milburn, 1992a) in which no normal trichromat made more than one error. The relationship of passing and failing to type and degree of color vision deficiency is shown in Table 2. All individuals with normal color vision passed both the CFPST and FPST. Only one individual with normal color vision made one error on the CFPST; normals made no errors on the FPST. The probability of failing the flight progress strips tests increased with degree of color vision deficiency. No dichromats, who have the greatest reduction in color discrimination ability, were able to pass either test. Only a few of the extreme anomalous category were able to pass both tests, and all were deutans. At least half of both protan and deutan simple anomalous trichromats were able to pass both tests. Overall, the impact of color vision deficiency on color discrimination in flight progress strips seems comparable in both protans and deutans in terms of pass/fail scoring (but not in terms of total errors as discussed below). Simple anomalous trichromats had approxi-

Table 1
Anomaloscope Classifications of Subjects

Normal Trichromat	Anomalous Trichromat: Simple Pro	Simple Deu	Extreme Pro	Extreme Deu	Dichromat Pro	Dichromat Deu
106	11	18	11	21	13	11

Table 2
Number of Subjects Passing and Failing on the Original FPS Test and the Copy FPS Test as a Function of Type and Degree of Color Vision Deficiency

		Normal Trichromat	Anomalous Trichromat: Simple Pro	Anomalous Trichromat: Simple Deu	Anomalous Trichromat: Extreme Pro	Anomalous Trichromat: Extreme Deu	Dichromat Pro	Dichromat Deu
CFPST	Pass	106	7	11	0	3	0	0
	Fail	0	4	7	11	18	13	11
FPST	Pass	106	7	9	0	5	0	0
	Fail	0	4	9	11	16	13	11

mately a 55% chance of passing, extreme anomalous trichromats had approximately a 16% chance of passing, and no dichromats were able to pass.

Error Scores. The mean and standard deviation of total error scores and 95% confidence intervals for the mean are shown for the CFPST and FPST in Table 3 as a function of type and degree of color vision deficiency. The confidence intervals are presented for comparison of means for color deficient groups and normal trichromats. The low, near zero variance among normal trichromats makes analysis of variance inappropriate for comparison of the means of color deficient groups and normals (Box, 1954). Error scores for both CFPST and FPST increase with degree of color vision deficiency, and the 95% confidence intervals for the means of all categories of deficients typically did not include the mean for normal trichromats; the only exception was the small overlap of the confidence interval for the simple protanomalous category on the FPST. Similarly, the mean number of errors for simple anomalous trichromats in both CFPST and FPST tasks is consistently below the 95% confidence interval for extreme anomalous and dichromat categories for both protans and deutans. Error scores are considerably higher than scores for the simple anomalous in both dichromats and extreme anomalous of both types. The mean error scores of the extreme anomalous groups were lower than the means for the dichromat groups, but within the 95% confidence intervals for dichromats for both FPST and CFPST error scores.

Analyses of variance, with normal trichromats excluded, were performed with the MANOVA program of the SPSS statistical system (SPSS, Inc., Chicago, IL). The error scores were evaluated with analysis of variance as a function of type and degree of color vision deficiency as "between groups" factors and test (CFPST versus FPST) as a "within groups" factor. The assumption of homogeneity of variance was rejected by the Box-M test. Since analysis of variance is robust for violation of this assumption when the number of subjects in each group is equal, the method of compensating for heterogeneity of variance by randomly discarding cases to achieve equal cell numbers was used (Tabachnick & Fidell, 1989; Toothaker, 1986). Findings of the "equal n" analysis are reported below. Each of 6 groups formed by the two types and three degrees of color vision deficiency had 11 cases. The main effects of both type and degree of color vision deficiency were statistically significant. Protans made significantly more errors than deutans on the FPS tests [$F(1,79)=97.52$, $p<.001$] and errors increased significantly with degree of color vision [$F(2,79)=70.88$, $p<.001$], as discussed above. The interaction of type and degree was also significant [$F(2,79)=28.91$,

Table 3
Errors on FPST and CFPST as a Function of Type and Degree of Color Vision Deficiency

	Diagnosis	Mean	Standard Deviation	Low	to	High
CFPST						
	Normal	0.01	0.10	0.01	to	0.03
	Protans					
	Simple	1.91	2.39	0.31	to	3.51
	Extreme	26.91	4.93	23.60	to	30.22
	Dichro.	29.54	7.58	24.96	to	34.12
	Deutans					
	Simple	3.50	5.83	0.60	to	6.40
	Extreme	7.67	4.20	5.76	to	9.58
	Dichro.	11.91	7.30	7.01	to	16.81
FPST						
	Normal	0	0	0	to	0
	Protans					
	Simple	2.73	4.94	-0.59	to	6.05
	Extreme	24.46	5.94	20.47	to	28.44
	Dichro.	27.92	9.12	22.42	to	33.43
	Deutans					
	Simple	2.94	5.35	0.28	to	5.61
	Extreme	5.48	5.07	3.17	to	7.78
	Dichro.	9.55	7.49	4.52	to	14.58

$p<.001$]; there was greater difference between protans and deutans among extreme anomalous trichromats and dichromats than among simple anomalous trichromats. There was also a significant difference for the main effect of test, the within groups factor [$F(1,79)=6.65$, $p=.012$]. No interaction involving the test factor was significant. The significant main effect of test reflects the fact that error scores tended to be slightly higher in the CFPST test than in the FPST. The dichromats and extreme anomalous groups averaged approximately two more errors on the FPST than the CFPST. Mean errors for the simple anomalous trichromats, however, were very similar in the two tests; the simple protanomalous actually had 0.82 more errors on the FPST while the simple deuteranomalous had 0.55 more errors on the CFPST. As shown below, any small difference in difficulty between the two tests was not reflected in pass/fail performance. The number passing both tests was exactly the same (127), as shown in the marginal totals of Table 4.

Validity of the FPST

The validity of the FPST was evaluated by comparing pass/fail performance on FPST to pass/fail performance on the CFPST. Cohen's Kappa (κ), an index of

Table 4
Pass/Fail Performance on the FPST versus CFPST

		FPST		
		Pass	**Fail**	**Total**
CFPST	Pass	123	4	127
	Fail	4	60	64
	Total	127	64	191
	Kappa=0.91			

agreement, was used to assess the validity of the practical test, as recommended by the National Research Council-National Academy of Science (NRC-NAS) Committee on Vision (1981). The index can be interpreted as the percentage agreement between test and criterion variable, with correction for chance. The relation of passing and failing on the FPST to passing and failing on the CFPST is shown in Table 4. The validity of the FPST for prediction of pass/fail performance on the CFPST was $\kappa = .91$. That validity is high. The miss *rate*, or probability of passing the FPST given that the CFPST was failed was .06. The *false alarm rate*, the probability of failing the FPST given that the CFPST was passed, was .03. Those error rates are considered acceptably low, and are comparable to error rates of the better clinical color vision tests when used to predict the criterion of normal versus abnormal classification by an anomaloscope test.

Reliability of the FPST

The FPSs of each test were split into two groups for computation of split-half reliability coefficients, one with odd and one with even numbered FPSs. The "raw" correlation between the two forms "halves" of the FPST was .91. The correlation between "halves" of the CFPST was .94. Since split-half coefficients can be affected by the particular selection of items that are assigned to each half, Chronbach's alpha was computed as a more conservative estimate of reliability. Chronbach's alpha is based on the average inter-correlation among the items of a test, or internal consistency of items, and the number of items (Nunnally, 1978). Since the "item" in this analysis was the error score for all items on an FPS, there were much larger error scores on some FPSs than others, thus yielding much different variances. Therefore, the "standardized item alpha" based on standardized item scores is cited for the FPST and CFPST. The alphas obtained were .93 for the FPST, and .95 for the CFPST. The alphas obtained for the six FPSs of the computer printing subscale were .91 for the FPST and .92 CFPST. The alphas obtained for the 21 FPSs of the handwriting subscale were .89 and .92 for the FPST and CFPST, respectively. It can be concluded that both FPST and CFPST tests and their subscales have high estimated reliability.

Aircraft Lights Test

Pass/Fail Performance. The ALT-L and ALT-H permit comparison of responses to the ALT when administered in a low illumination similar to the illumination at night in an ATC tower, and a relatively high illumination characteristic of an office. The number of subjects passing and failing on the ALT are shown in Table 5 as a function of type and degree of color vision deficiency. The criterion for passing the ALT-L was no more than one error; the criterion for the ALT-H was the same. All individuals with normal color vision passed both ALT-L and ALT-H. Only one error was made by a normal in an ALT test and that was on the ALT-L. The probability of failing the ALT increased with degree of color vision deficiency. No dichromats or extreme anomalous trichromats were able to pass the ALT, with the exception of one extreme deutan who passed the ALT-H test. The proportions passing among simple protans and simple deutans were less than 50%.

Table 5

Number of Subjects Passing and Failing on the ASL-L Test and the ASL-H Test as a Function of Type and Degree of Color Vision Deficiency

		Normal Trichromat	Anomalous Trichromat Simple Pro	Simple Deu	Extreme Pro	Extreme Deu	Dichromat Pro	Dichromat Deu
ASL-L	Pass	106	4	7	0	0	0	0
	Fail	0	7	11	11	21	13	11
ASL-H	Pass	106	3	9	0	1	0	0
	Fail	0	8	9	11	20	13	11

Table 6

Errors on ASL-L and ASL-H as a Function of Type and Degree of Color Vision Deficiency

	Diagnosis	Mean	Standard Deviation	Low	to	High
ASL-L						
	Normal	0.01	0.10	-0.01	to	0.03
	Protans					
	Simple	4.00	5.12	0.56	to	7.44
	Extreme	18.27	4.22	15.44	to	21.11
	Dichro.	17.69	4.44	15.01	to	20.38
	Deutans					
	Simple	6.39	7.13	2.84	to	9.94
	Extreme	16.48	6.88	13.34	to	19.61
	Dichro.	19.00	6.02	14.96	to	23.04
ASL-H						
	Normal	0	0	0	to	0
	Protans					
	Simple	6.00	5.02	2.63	to	9.37
	Extreme	19.55	2.65	17.76	to	21.33
	Dichro.	17.54	4.59	14.76	to	20.31
	Deutans					
	Simple	6.78	7.99	2.81	to	10.75
	Extreme	16.95	6.01	14.22	to	19.69
	Dichro.	18.00	4.92	14.70	to	21.31

Error Scores for the ALT. The mean and standard deviation of error scores and 95% confidence intervals for the mean are shown in Table 6 as a function of type and degree of color vision deficiency. Errors increased with degree of deficiency and the performance of protans and deutans was similar. The mean of normals was zero or near zero for both tests. The 95% confidence intervals of all color deficiency categories did not include the mean of normals, and the means for both simple anomalous groups were below the 95% confidence limits for the extreme anomalous and dichromat groups. Means for the extreme anomalous were slightly higher than the means

for dichromats among protans and lower than dichromats among deutans, but in both cases means were not outside the limits of the other group.

The ALT error scores for color deficients were evaluated by analysis of variance as a function of type and degree of color vision deficiency as "between groups" factors and test, ALT-L versus ALT-H, as a "within groups" factor; normal trichromats were excluded. The same cases that were dropped in the FPS analysis of variance were again dropped in this analysis to achieve an equal number of subjects (11) in each group in order to compensate for heterogeneity of variance. Only the main effect of degree of deficiency was significant in this analysis [$F(2,60)=41.15$, $p<.001$]. The main effects of type of deficiency and test were not significant, nor were any interactions. This result suggests that ALT performance in individuals with varying type and degree of color vision deficiency is relatively

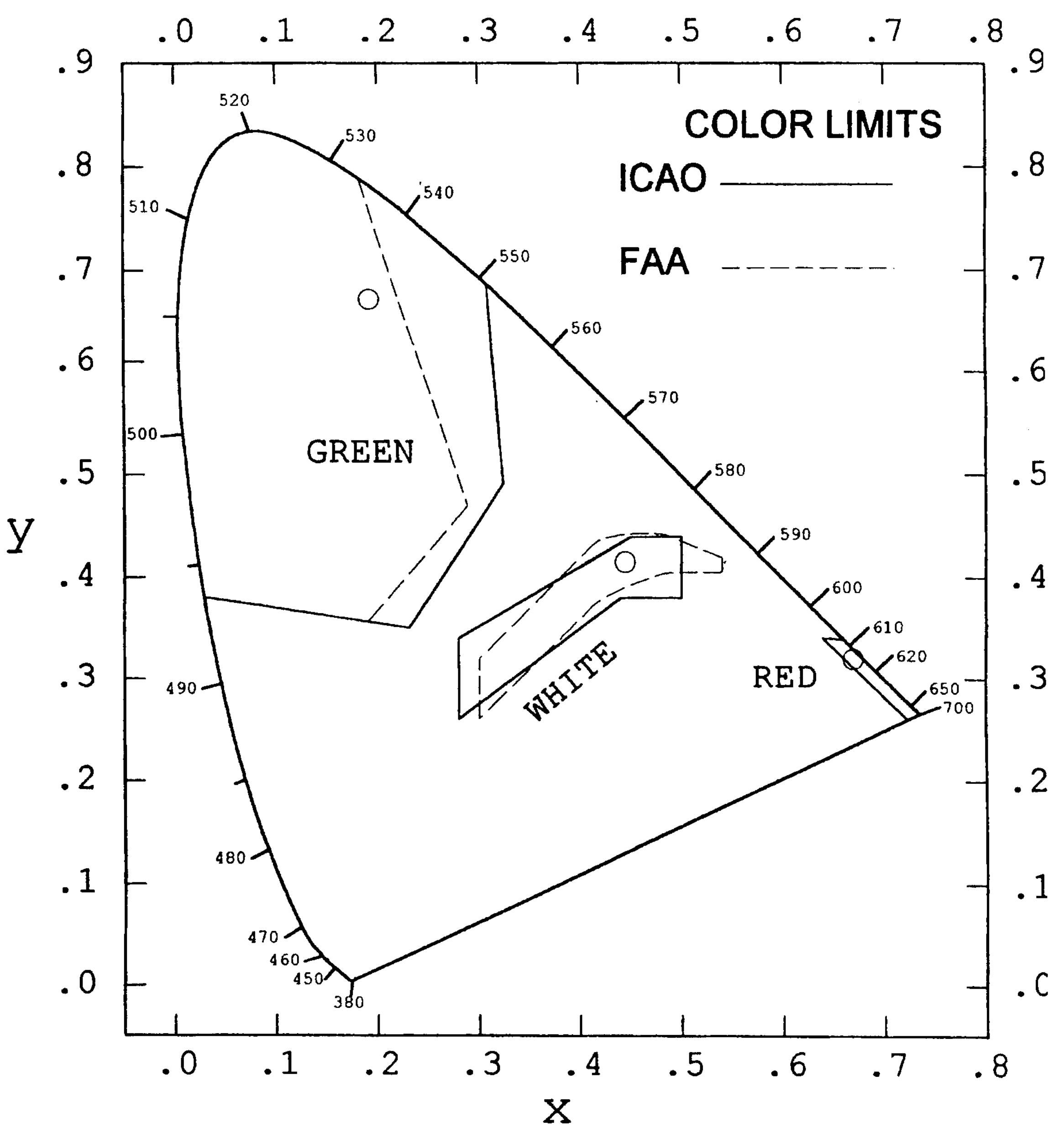

Figure 1. Chromaticity diagram (CIE 1931) showing chromaticity of lights of the Aviation Lights Test and the color limits for aviation signal colors as specified by the FAA and ICAO.

insensitive to variation in test room illumination. The relatively minor variations in test room illumination normally occurring as a function of room size, lightness of walls and so forth, should not affect performance of individuals taking the test.

Validity of the ALT. The light point stimuli of the ALT were validated physically with the computer controlled spectroradiometer mentioned above. The measured chromaticities of test lights met the physical color specifications for aviation signal red, green, and white lighting as given in ICAO regulations and FARs. The measured chromaticities of the red, green, and white lights of the ALT are given in Figure 1, and are plotted along with the gamuts for the three signal colors as specified by the ICAO.

Reliability of the Practical Color Vision Tests. The raw correlation between halves was .96 for both the ALT-L and ALT-H tests. Reliabilities, estimated by Chronbach's alpha, were .98 and .97 for ALT-L and ALT-H tests, respectively. The three series of nine light point pairs that were presented in each test were considered subscales; alphas were .92, .93, and .92 for the three subscales of the ALT-L. The alphas for the three subscales of the ALT-H were .93, .92, and .92. In summary, the reliabilities of the ALT tests were very high.

CONCLUSIONS

The present research is the third study in a series of studies on the ATCS color vision standard that addressed three problems (1) whether a color vision standard was needed (Mertens & Milburn, 1992a), (2) whether FAA-accepted clinical color vision tests were valid for screening ATCS applicants (Mertens & Milburn, 1992b), and (3) the development, evaluation, and validation of two practical color vision tests for the two largest options within ATC work, the en route center and terminal options. While these studies have shown that error in simulated ATCS color tasks is rare in normal trichromats and that many individuals with color vision deficiency are prone to error, the research has also shown that there are some individuals with color vision deficiency that can perform the tasks as well as normals. The research evaluating the validity of clinical tests has shown that the validity of some tests is high and that the error rate is generally acceptable in terms of not passing individuals who cannot perform ATCS color tasks accurately. Those tests have higher error rates in terms of failing some individuals who can perform ATCS color tasks as accurately as normals. The present research has shown that practical, job-sample color vision tests can be constructed that have high face validity and also high demonstrated validity in terms of selecting individuals who can perform ATCS color tasks as accurately as normals, and that also have very high reliability as required of good selection tests. These two practical color vision tests yield effective and fair color vision screening of job-relevant color vision ability for ATCS applicants, and have been in use for more than one year.

REFERENCES

Adams, A. J. & Tague, M. K. (1985). Performance of air traffic control tasks by protanopic color defectives. *American Journal of Optometry and Physiological Optics*, 62, 744-50.

Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems. I. Effect of inequality of variance in the one-way classification. *Annals of Mathematical Statistics*, 25, 290-302.

Dain, S. J., Honson, V., & Ang, J. (1988). The effect of two lighting conditions on performance of the Farnsworth lantern color vision test. *Aviation, Space, and Environmental Medicine*, 59, 371-3.

Federal Aviation Administration. (1989). Air traffic control. (FAA Order 7110.65F). Washington, D. C.: Author.

Jewell, L. N., & Siegall, M. (1990). *Contemporary industrial and organizational psychology.* (2nd Ed.). St. Paul, MN: West Publishing Co.

Kaufman, J. E., & Christensen, J. F. (Eds.). (1972). *IES lighting handbook* (5th Ed.). New York, Illuminating Engineering Society of North America. (p. 21-15)

Kaufman, J. E., & Christensen, J. F. (Eds.). (1987). *IES lighting handbook: Applications Volume.* New York: Illuminating Engineering Society of North America. (p. 15-2)

Kuyk, T. K., Veres III, J. G., Lahey, M. A., & Clark, D. J. (1986). The ability of protan color defectives to perform color-dependent air traffic control tasks. *American Journal of Optometry and Physiological Optics*, 63, 682-6.

Kuyk, T. K., Veres III, J. G., Lahey, M. A., & Clark, D J. (1987). Ability of deutan color defectives to perform simulated air traffic control tasks. *American Journal of Optometry and Physiological Optics*, 63, 682-6.

Lahey, M. A., Veres III, J. G., Kuyk, T. K., Clark, D. J., & Smith, E. N. (1984). Job analysis and determination of color vision requirements for air traffic control specialists. Report submitted under Contract OPM-50-83 with the U.S. Office of Personnel Management.

Mertens, H. W., & Milburn, N. J. (1990). Evaluation of functional color vision requirements and current color vision screening tests for air traffic control specialists. Report DOT/FAA/AM-90/9. Washington, D. C.: Federal Aviation Administration, Office of Aviation Medicine.

Mertens, H. W., & Milburn, N. J. (1992). Performance of color-dependent tasks of air traffic control specialists as a function of type and degree of color vision deficiency. Report DOT/FAA/AM-92/28. Washington, D. C.: Federal Aviation Administration, Office of Aviation Medicine. (a)

Mertens, H. W., & Milburn, N. J. (1992). Validity of clinical color vision tests for air traffic control specialists. Report DOT/FAA/AM-92/29. Washington, D. C.: Federal Aviation Administration, Office of Aviation Medicine. (b)

National Research Council-National Academy of Sciences, Committee on Vision. (1981). *Procedures for testing color vision: Report of Working Group 41.* Washington, D. C.: National Academy Press.

Nunnally, J. C. (1978). *Psychometric theory.* New York: McGraw-Hill. (p. 229)

Steen, J. A., Collins, W. E., Lewis, M. F. (1974). Utility of several clinical tests of color defective vision in predicting daytime and nighttime performance with the aviation signal light gun. *Aerospace Medicine*, 45, 467-72.

Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics*, New York: Harper and Row. (p. 48-49)

Toothaker. L. E. (1986). *Introductory statistics for the behavioral sciences.* New York, McGraw-Hill. (p. 452)

U. S. Office of Personnel Management. (July 1993). *Qualifications Standards Handbook for General Schedule Positions* (Replaces Handbook X-118), Washington, D. C.: Author.

Vortac, O. U., Manning, C. A., and Rotter, A. J. (1992) En route air traffic controller's use of flight progress strips: A graph-theoretic analysis. Report DOT/FAA/AM-92/31. Washington, D. C.: Federal Aviation Administration, Office of Aviation Medicine.

☆U.S. GOVERNMENT PRINTING OFFICE: 1995 - 660-067/00128